

CENTRE FOR  
VOCATIONAL  
EDUCATION  
RESEARCH



# The long-term impact of improving non-cognitive skills of adolescents: Evidence from an English remediation programme

Augustin de Coulon, Vahé Nafilyan, Stefan Speckesser

Research Discussion Paper 028

March 2020

The Centre for Vocational Education Research (CVER) is an independent research centre funded by the UK Department for Education (DfE). CVER brings together four partners: the LSE Centre for Economic Performance; University of Sheffield; National Institute of Economic and Social Research and London Economics.

Any views expressed are those of the authors, and do not represent the views of DfE. For more details on the Centre, go to [cver.lse.ac.uk](http://cver.lse.ac.uk)

Published by:  
Centre for Vocational Educational Research  
London School of Economics & Political Science  
Houghton Street  
London WC2A 2AE

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means without the prior permission in writing of the publisher nor be issued to the public or circulated in any form other than that in which it is published.

Requests for permission to reproduce any article or part of the Working Paper should be sent to the editor at the above address.

© A. De Coulon, V. Nafilyan, S. Speckesser, March 2020.

# The long-term impact of improving non-cognitive skills of adolescents: Evidence from an English remediation programme

Augustin de Coulon<sup>1,2</sup>, Vahé Nafilyan<sup>1,2,3</sup>, and Stefan Speckesser<sup>4,2</sup>

<sup>1</sup>King's Business School, King's College London

<sup>2</sup>Centre for Vocational Education Research

<sup>3</sup>Office for National Statistics

<sup>4</sup>National Institute of Economic and Social Research

March 11, 2020

**Abstract:** In this paper, we estimate the long-term labour market effects of Entry to Employment (E2E), an intervention designed to improve the non-cognitive skills of low-achieving adolescents in England. Using an instrumental variable (IV) approach, we find that E2E courses substantially increased earnings of participants in the long-run. The increase is primarily driven by a large and significant effect on the probability to be in employment. Placebo tests and robustness checks provide further support that the link is unlikely to be affected by unobserved confounders.

**Acknowledgements:** We thank Prof. Sandra McNally and participants to the 2019 Centre for Vocational Education conference for helpful comments. We would like to thank the data sharing team of the Department for Education, HM Government, for their help with data access (data share agreement DR160321.03). This work has been funded by the UK Department for Education.

**Corresponding author:** Vahé Nafilyan, King's Business School, King's College London, Strand, London WC2R 2LS, United Kingdom, Email: vahe.nafilyan@kcl.ac.uk

# Executive summary

## Aim

Designing programmes to help low achieving young people make successful school-to-work transitions is notoriously difficult. Whilst the positive effects of programmes targeted towards infants and young children are well-documented (Heckman et al. 2010, Chetty et al. 2011), there are doubts about the effectiveness of remediation programmes targeted towards adolescents (Kautz et al. 2014). But few programmes targeting adolescents include long-term follow-ups, and therefore little is known about the long-term effects of such interventions.

In this paper, we estimate the long-term labour market effects of a remedial intervention, the now inactive *Entry to Employment* programme (E2E). E2E aimed to improve labour market outcomes of low-achieving adolescents in England by targeting their non-cognitive skills.

## Background

The importance of non-cognitive and soft skills for successful transitions into the labour market has been long established in the economics literature (Borghans, Duckworth, Heckman & ter Weel 2008, Almlund et al. 2011). Yet, the evidence about the effectiveness of interventions aiming to improve the non-cognitive skills amongst disadvantaged adolescents is mixed (Schochet et al. 2008, Holmlund & Silva 2014, Martins 2017, Heller et al. 2017). The newly available Longitudinal Education Outcomes (LEO) data offers new opportunities to understand the longer-term impact of such programmes, i.e. ten years after participation or later.

We combine data on participants in the Entry to Employment (E2E) programme to pre-programme school achievement data and long-term employment and earnings records retrieved from income tax registers. Running between 2003 and 2010, E2E was tailored to individual needs but primarily focused on non-cognitive skills with the aim to improve young people's transitions to the labour market or training. Unlike the majority of education programmes, E2E also did not lead to recognised qualifications.

## **Identification strategy**

Participation in E2E programme is endogenous. Summary statistics suggest that young people who enrol in E2E have lower prior attainment and are more likely to come from poor families compared to those who did not enrol in any post-16 education. To overcome the endogeneity issue, we exploit the variation in local availability of the programme as an instrument for E2E participation: E2E was not available in 37 out of the 326 Local Authority Districts (LADs), while in other LADs over a third of all providers offered E2E courses.

The potential issue with this approach is that young people living in LADs where no or only a few providers offer E2E could differ along many characteristics. Summary statistics indicate that they are less likely to come from disadvantaged families or have better prior attainment compared to young people living in LADs that offer E2E. To mitigate this issue, we include in our model a range of local area-level characteristics as well as a range of socio-demographic characteristics, including prior attainment at school exams. Nonetheless, it is possible that living in LADs that do not offer E2E courses is associated with unobserved characteristics.

To test whether the instrument is capturing the effect of unobserved characteristics, we conduct two set of placebo tests. First, we estimate the reduced-form effect of the instrument on the pre-determined educational outcomes (results at KS4, KS3 and KS2) of both E2E participants and those who left school at 16, conditioning on observed area level characteristics. Second, we estimate the association of the instrument with the labour market outcomes of those who studied for A-levels after leaving school. The proportion of providers that offer the E2E programme in the home LAD should not be associated with pre-determined educational outcomes nor the labour market outcomes of those who studied for A-levels. A statistically significant association would indicate that the instrument acts as a proxy for unobserved area characteristics, and therefore that is not valid.

## **Findings**

We find that participating in E2E courses substantially increased the earnings of the participants in the medium and long-run. The IV estimate suggests that E2E increased earnings by over £7,000 10 to 12 years after participating in the programme. The increase in earnings is primarily

driven by an increase in the probability to be employed. The results from our placebo tests indicate that the instrument is valid and does not act as a proxy for unobserved confounding factors.

We find significant long-term benefits from participation in the programme. The benefits of the programme took several years to materialise, implying that an early evaluation of the programme would have led to the erroneous conclusion that the programme had no or little effect on labour market outcomes. Our paper demonstrates the importance of focusing on long-term outcomes when evaluating programmes targeting young people.

# 1 Introduction

Designing programmes to help low achieving young people make successful school-to-work transitions is notoriously difficult. Adolescent remediation programmes are not as effective as programmes targeting infants and young children (Kautz et al. 2014). Because early childhood is a critical period for formation of both cognitive and non-cognitive skills (Heckman 2006), and the dynamic nature of skill formation, early childhood programmes can have long-lasting positive effects on many dimensions of life (Heckman et al. 2010, Chetty et al. 2011). By contrast, there are doubts about the effectiveness of remediation programmes aiming to improve the cognitive skills of adolescents. The evidence on the effect of remedial education programmes offers mixed results (Jacob & Lefgren 2004, Lavy & Schlosser 2005, Rodríguez-Planas 2012, Taylor 2014).

Designing programmes to help low achieving young people make successful school-to-work transitions is notoriously difficult. Adolescent remediation programmes are not as effective as programmes targeting infants and young children (Kautz et al. 2014). Because early childhood is a critical period for formation of both cognitive and non-cognitive skills (Heckman 2006), and the dynamic nature of skill formation, early childhood programmes can have long-lasting positive effects on many dimensions of life (Heckman et al. 2010, Chetty et al. 2011). By contrast, there are doubts about the effectiveness of remediation programmes aiming to improve the cognitive skills of adolescents. The evidence on the effect of remedial education programmes is disparate (Jacob & Lefgren 2004, Lavy & Schlosser 2005, Rodríguez-Planas 2012, Taylor 2014).

Whilst cognitive skills become solidified around the time of puberty, non-cognitive skills remain malleable until later ages, due to the slower development of the prefrontal cortex. As a result, remedial interventions targeted at adolescents focusing on non-cognitive skills are expected to be the most effective (Kautz et al. 2014). The importance of non-cognitive and soft skills for successful transitions into the labour market has been long established (Borghans, Duckworth, Heckman & ter Weel 2008, Borghans, Weel & Weinberg 2008, Almlund et al. 2011, Heckman & Kautz 2012).<sup>1</sup> Yet, the evidence about the effectiveness of interventions

---

<sup>1</sup>Soft skills are associated with higher educational achievement (Carneiro et al. 2007), increased productivity (Feinstein 2000, Machin et al. 2001), lower absenteeism (Strmer & Fahr 2013), better work relationships (Caliendo et al. 2015), shorter unemployment spells (Lindqvist & Vestman 2011, Gallo et al. 2003) and compensation scheme choice (Dur et al. 2010).

aiming to improve the non-cognitive skills amongst disadvantaged adolescents is mixed. For instance, Job Corps, the US largest training program for disadvantaged youths increases educational attainment and reduces criminal activity but has no long-term effects on earnings (Schochet et al. 2008). Evaluating an education intervention targeting underachieving pupils non-cognitive skills in England, Holmlund & Silva (2014) find that it reduced absenteeism, but failed to improve cognitive skills. By contrast, Martins (2017) finds that a remediation programme targeting low achieving students' non-cognitive skills in Portugal improved educational achievement. A mentorship programme that targets non-cognitive skills of disadvantaged youth in Chicago ('Becoming a Man') increased grades (Cook et al. 2014) and reduced criminal activity (Heller et al. 2017). Very few programmes targeting adolescents have long-term follow-ups, and therefore little is known about the medium and long-term effects of these interventions.

In this paper, we evaluate the labour market effects of Entry to Employment (E2E), a remedial intervention designed to improve the school-to-work transition of low-achieving and disadvantaged young people in England, which was introduced nationally in 2003. E2E aimed to equip young people with the skills needed to start an apprenticeship, other vocational courses or find a job. The learning delivered by the E2E programme was tailored to the participants needs but primarily focused on non-cognitive skills.

We estimate the medium and long-term effects of this programme on individual earnings and employment outcomes based on data for a cohort of school leavers in England in 2003. As a control group, we use the young people from the same cohort who did not participate in any post-16 education<sup>2</sup>.

First, we compare the labour market outcomes of E2E participants to those without further education. Although E2E participants have worse labour market outcomes than this comparison group, we find that participation in E2E is associated with an increase in annual earnings once we adjust for a set of socio-demographic characteristics and prior educational attainment.

To account further for the endogeneity of participation in E2E, we use the availability of E2E courses in the home Local Authority District (LAD) as an instrument for participation in E2E. There is substantial heterogeneity in the availability of E2E courses. In some LADs, this

---

<sup>2</sup>An approach commonly used in this literature, see Wolter and Ryan, 2011, Adda et al. 2014)

programme is offered by more than a third of all providers; In more than 10 per cent of LADs (37 out of 326), mainly the more affluent ones, E2E is not offered by any providers. Even when controlling for LAD and individual characteristics, availability of E2E is a strong predictor of actual programme participation

Our results show that the programme generated substantial benefits in the medium and long-term. The Instrumental Variables (IV) estimate suggests that E2E participation increased earnings by over £9,000 around 10 years after the programme participation. We conduct two placebo tests to confirm the validity of our instrument. First, we show that the instrument is not associated with educational attainment at age 11, 14 and 16, once we adjust for area and socio-demographic characteristics. Second, we show that there is no association between the instrument and the labour market outcomes of young people who studied towards A-Levels, the end of upper secondary school qualifications in the UK. This indicates the instrument does not act as a proxy for unobserved area characteristics.

We contribute to the existing literature by showing that remedial interventions aiming to improve non-cognitive skills can substantially improve the labour market outcomes of low-achieving adolescents. We also demonstrate the importance of focusing on long-term outcomes. Focusing solely on short-term benefits would have led to the erroneous conclusion that the programme had no or little effect on labour market outcomes.

In this paper, we first describe the E2E programme. We then present the data and show some summary statistics for our treatment and control group. Third, we describe our identification strategy and then present and discuss our results and sensitivity analyses.

## **2 Entry to Employment**

The E2E programme was a post-16 education programme introduced in the 2002/03 academic year in selected pathfinder areas and rolled out nationally in the subsequent year. Its aim was to offer support for young people 'not yet ready or able to take up a Modern Apprenticeship, an intermediate-level vocational education or training programme or move directly into employment' (Spielhofer et al. 2003). It replaced a programme focusing on Life Skills, Preparatory Training and Other Training after the Cassels Report (Department for Education and Skills

2001) suggested that a pre-employment programme should be set up for people not yet able to engage in apprenticeships or employment. This recommendation was taken forward by an HM Government Green Paper (14-19: Extending Opportunities, Raising Standards 2002). It targeted young people with very low attainment, few qualifications and low on self-confidence. It included those with learning difficulties and disabilities. The E2E aimed to reduce the number of people leaving the education system to become inactive (i.e. Not in Education, Employment and Training, NEET). While the programme involved some English and Maths, the topics were often embedded in a workplace context (e.g. reading and understanding of work instructions). A large part of E2E teaching was directed at raising personal and social skills, confidence building, managing own feelings and working in a team.

Vocational skills were included through work-experience opportunities, understanding how career aspirations relate to qualifications and general employability skills (e.g. timekeeping, understanding business objectives). The structure of the course varied according to the specific needs identified by one dedicated caseworker, who advised and delivered mentoring (Spielhofer et al. 2003). The programme involved recognised qualifications only when deemed appropriate.

While precise figures on enrolment into 'Entry to Employment' are not available in official statistics, the programme was expected to be large when operating at steady state. In the early stages, up to 50,000 learners were expected to participate in the programme each year (Select Committee on Education and Skills 2003). Spending was around 170 million pounds per year (Learning and Skills Council 2009), about a quarter of what was spent on young people apprenticeships (630 million).

A qualitative evaluation study accompanying the trial phase of the programme found that E2E participants had very low expectations at programme entry but that those had increased markedly at completion. E2E was seen as unsuccessful at the time as only one out of two people going through the programme reported gainful employment (Wolf 2011). Relatively few participants became employed immediately after the programme (Deere 2016). Seen as widely unsuccessful based on those and without evidence of potential long-term benefits, E2E was replaced by a new programme Foundation Learning in 2010, which (unlike E2E) included the requirement of formal and recognised qualifications.

## 3 Data and summary statistics

### 3.1 Data

We examine the labour market outcomes of the young people who left secondary school in the summer of 2003 and participated in the Entry to Employment (E2E) programme at any time during the following three academic years (2003/06 to 2005/06). The data combine individual-level records from National Pupil Data (NPD) and the Individualised Learner Records (ILR) on participation in general and vocational education after the age of sixteen.

The NPD holds a wide range of information about all students who attend publicly-funded schools and colleges in England. The NPD combines the examination results of pupils with information on pupil and school characteristics. It contains a range of socio-demographic characteristics, including gender, ethnicity, special educational need status (e.g. disability, learning difficulties), eligibility to Free School Meal. It also includes results from national examinations taken at the age of 11 (Key Stage 2), 14 (Key Stage 3), 16 (Key Stage 4) and 18 (Key Stage 5).

The ILR collects data about learners in the Further education sector. It contains information on every learning course undertaken in England by a publicly-funded provider, such as course level, duration, and outcome. It also includes information on learners, such as key demographic characteristics. We use these data to identify young people who participated in the E2E programme. Among the 2003 school leavers, we identify 34,758 young people who participated in E2E between 2003/04 to 2005/06, about 5 per cent of all.

The E2E programme targeted young people at risk of leaving the educational system and become unemployed or inactive. To obtain a potential comparison group, we identify 64,502 people from this cohort who did not participate in any post-16 education in the same time period, i.e. people not having any further learning recorded in NPD or ILR data. We use these young people as a control group because they are similar to those who were targeted by the programme. In our sensitivity analyses, we show that we obtain similar results using alternative control groups.<sup>3</sup>

---

<sup>3</sup>In one specification, we restrict the control group to those who did not participate in any post-16 education in the three academic years after leaving school and were not employed in the 2004 tax year. In another, we use young people who engaged in low-level vocational courses other than E2E as a control group.

As measures of deprivation at the local area level, we use the Income Deprivation Affecting Children Index (IDACI). This measures the proportion of all children aged 0 to 15 living in income deprived families at the Local Super Output Area (LSOA) level. The LSOA is a small geographical area, with a population of around 1,500 on average. England and Wales are divided into 34,753 LSOA. We further add the proportion of working age adults claiming unemployment benefits in 2003 at the Local Authority District level, based on data from the Department for Work and Pension<sup>4</sup>. As a measure of the educational profile of the local area, we use the proportion of adults holding qualifications of various levels for each LSOA, based on data from the 2001 Census.<sup>5</sup>

Data on labour market outcomes were obtained from Her Majesty's Revenue and Customs (HMRC) records, including information on earnings for each tax year between 2004/05 and 2015/16. The HMRC data cover all employees who earn more than the tax-free allowance but exclude the self-employed. We retrieved HMRC tax records for 93.6% of E2E participants and 75.4% of those without post-16 education. The main reason why these individuals were not retrieved from the HMRC tax records is that they may never have been in formal paid employment in the UK. They may also not have been matched because of incomplete or missing key identifiers, but this affects only a small proportion in linked administrative data (usually less than 5%). As a result, we consider that those not retrieved in the HMRC tax records have no earnings. In an alternative specification, we restrict our sample on those who were retrieved in HMRC data. We find that this restriction does not affect our results.

We also use complete ILR data to measure the availability of provision of E2E courses at the Local Authority District (LAD) level. We compute the proportion of providers that offer E2E courses in the 2003/04 school year at the LAD level.<sup>6</sup>

---

<sup>4</sup>These data were retrieved from NOMIS

<sup>5</sup>It includes the proportion of adults holding no qualifications, qualifications of level 1, level 2, level 3 or level 4 or above. These data were also retrieved from NOMIS

<sup>6</sup>Due to missing postcodes information in ILR, we could not estimate the proportion for the following LADs: East Hertfordshire, Gateshead, Northumberland, St Albans, Stevenage, Welwyn Hatfield. We exclude from our analytical sample the 1,602 young people who lived in these LADs whilst completing their last year in secondary school.

## 3.2 Descriptive evidence

Summary statistics for the sample are presented in Table 1. The first column shows the mean of individual characteristics and outcomes for the young people who did not participate in any post-16 education between 2003/04 to 2005/06. In column 2 we show the mean for those who participated in E2E courses between 2003/04 to 2005/06. The difference between the two is reported in column 3.

We find that E2E participants differ from those without post-16 education across almost all characteristics except for first-language. About 92 per cent have English as their first language, which is similar to those who did not take part in any post-16 education. Compared to those who did not participate in any further education, E2E learners are more likely to be male, have special educational needs (44 compared to 28 per cent), were eligible to Free School Meals during secondary school (32 per cent compared to 22 per cent) and to live in a more deprived area (as measured by the IDACI). They also performed less well at school, measured by the number of General Certificates of Secondary Education (GCSEs) they obtained. On average, participants have 0.6 GCSEs graded A\*-C, the grade that would allow them to progress on the academic track, and 4.9 GCSE at any grade, compared to 1.6 and 6.1 respectively for those who left the educational system altogether. Similarly, they were less likely to have achieved the expected level at examinations taken at age 14 (Key Stage 3 - KS3) and at examinations taken at age 11 (Key Stage 2 - KS2).

Table 1: Descriptive Statistics

	No E2E	E2E	Diff.	Obs.
<i>Demographic characteristics</i>				
Female	0.42	0.39	0.03***	99260
Special educational needs	0.28	0.44	-0.16***	99260
English as first language	0.92	0.92	0.00*	99260
White	0.85	0.83	0.02***	99260
Free school meal	0.22	0.32	-0.10***	99260
<i>Prior educational attainment</i>				
Number of A*-C GCSE	1.51	0.65	0.87***	99260
Number of GCSEs achieved	6.01	5.01	1.00***	99260
GCSE A*-C in English	0.46	0.28	0.19***	99260
GCSE A*-C in Maths	0.30	0.15	0.15***	99260
GCSE A*-C in Science	0.31	0.17	0.14***	99260
KS3 mat Level 5 or above	0.43	0.27	0.16***	99260
KS3 eng Level 5 or above	0.39	0.23	0.16***	99260
KS3 sci Level 5 or above	0.38	0.24	0.14***	99260
KS2 eng Level 4 or above	0.51	0.35	0.16***	99260
KS2 mat Level 4 or above	0.52	0.36	0.15***	99260
KS2 sci Level 4 or above	0.59	0.43	0.15***	99260
<i>Area characteristics</i>				
IDACI decile	4.39	3.60	0.79***	99260
Local UB claimant rate	2.37	2.67	-0.31***	99260
<i>Proportion of adults in the LSOA with:</i>				
No qualifications	0.34	0.37	-0.03***	99260
Level 1 qualifications	0.18	0.18	0.00***	99260
Level 2 qualifications	0.19	0.18	0.01***	99260
Level 3 qualifications	0.07	0.07	0.00***	99260
Level 4/5 qualifications	0.15	0.13	0.02***	99260
Other qualifications/level unknown	0.07	0.07	0.00***	99260
<i>Outcomes</i>				
Level 2+ qual.	0.06	0.33	-0.27***	99260
Pay in 2015 (incl 0)	8,570.6	7,474.9	1,095.7***	99260
Employed in 2015	0.56	0.65	-0.09***	99260
Days employed in 2015	197.08	223.74	-26.66***	99260
Days employed in 2015 - excl 0	350.13	344.47	5.67***	58882
Pay in 2015 (excl 0)	15,522.1	12,293.0	3,229.2***	56750
<i>Availability of E2E</i>				
Proportion of providers providing E2E	0.20	0.22	-0.03***	99260
Number of FE providers	19.38	21.56	-2.18***	99260

Note: Mean outcomes for E2E participants in the three year after leaving school and those with no post-16 education in the same period. Number of E2E participants: 34,758; Number of people who did not participate in any post-16 education: 64,502. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1 Source: NPD, ILR and HMRC data, KS4 leavers 2002/03

Based on ILR and KS5 data we identify the level of the highest qualification held by the individuals in our sample. These qualifications can have been obtained at any time between 2003/04 and 2014/15. We find that a third of those who started E2E achieve an intermediate vocational or general qualification ("Level 2") or higher by age 27/28. five times more than

those with no post-16 education in the first three years after leaving school.

Using data from HMRC we observe earnings and number of days employed in all tax years from 2004/05 to 2015/16. The average annual earnings in the 2015/16 tax year, including those who did not work, was £7,475 for E2E participants. Nearly two-thirds of E2E participants were employed in the tax year. They were employed on average for 344 days in the 2015 year and earned on average £12,293. Those who did not participate in any post-16 education have higher average earnings (£8,571 in 2015) than E2E participants but are less likely to be in employment (56 per cent). Those who are employed earn more than E2E participants: they earn on average £15,522, £3,230 more than E2E participants who were in paid employment. They also spent on average more days in employment (350).

Figure 1: Labour market outcomes of E2E participants and control group



Figure 1 displays average labour market outcomes for all tax years between 2004/05 to 2015/16 for those who participated in E2E and those who left the educational system at 16. We see that E2E learners have on average lower earnings than those who left school at 16 in every tax year. This holds regardless of whether we include and exclude those not employed. They

are more likely to be employed in any of the tax years (defined as having done at least one day of paid work during the year), but those who are employed worked on average fewer days than those who left school at 16.

## 4 Empirical Approach

The descriptive statistics reported in Table 1 above suggest that young people who participate in E2E courses have a substantially different background from those who left the education system after the end of secondary school. For instance, they come from more disadvantaged families and have poorer educational attainment. To estimate the effect of the E2E programme, the empirical design needs to address the issue that participation in E2E is not random. Our first approach is to estimate the effect of E2E participation on labour market outcomes by comparing the labour market outcomes of E2E participants to those of the control group, while adjusting for observed characteristics. More formally we estimate the following model:

$$y_i = \beta E2E_i + x_i\gamma + \varepsilon_i \quad (1)$$

Where  $y_i$  is our outcome of interest for individual  $i$ ;  $E2E_i$  is a binary variable indicating if the individual participated in a E2E course.  $x_i$  contains a range of individual and area characteristics that are likely to be correlated with both E2E participation and labour market outcomes. Area characteristics include region fixed effects, the deciles of the Income Deprivation Affecting Children Index (IDACI), the proportion of adults holding qualifications of various levels for each LSOA, and a measure of unemployment in September 2003 (the number of claimants divided by working-age population) at LAD level<sup>7</sup>. Individual characteristics include demographics (gender, special educational needs, ethnicity and whether English is their first language) and educational attainment measured by the number of GCSEs graded A\*-C. We also control for results to tests in English, Mathematics and Science administered at the age of 14 (KS3) and 11 (KS2).

Equation 1 relies on the assumption that that each control variable is linearly and indepen-

---

<sup>7</sup>The Income Deprivation Affecting Children Index measures the proportion of all children aged 0 to 15 living in income deprived families

dently related to the outcome. To test for potential non-linearity and interactions between these variables, we estimate a partially linear model in which the relationship between the control variables  $x_i$  and the outcome is unspecified:

$$y_i = \beta E2E_i + g(x_i) + \varepsilon_i \quad (2)$$

where  $g(\cdot)$  is an unknown function which can be estimated using a range of models. We use the Double Debiased Machine Learning (DML) estimator proposed by Chernozhukov et al. (2018) to estimate  $\beta$  and derive confidence intervals. We model separately as a function of the control vector ( $x_i$ ) the outcome and the participation in E2E. We then use the Generalised Method of Moments (GMM) to combine the residuals and obtain an estimate of  $\beta$  and derive confidence intervals.

We derive estimates of  $\beta$  using two different modelling approaches. First, we interact all characteristics from  $x_i$  and use the Least Absolute Shrinkage and Selection Operator (LASSO) to estimate the model. Second, we use gradient boosted trees to estimate the model in a non-parametric way (Chen & Guestrin 2016). In both approaches, the model is estimated in one part of the data, and the other part is used to produce the estimator using the GMM. To avoid losing efficiency, we use cross-validation, with 5 partitions.<sup>8</sup>

The shortcoming of both the linear and partially linear models is that they rely on the assumption that E2E learners and those who did not participate in any FE courses do not differ in terms of unobserved characteristics, once we adjust for observed heterogeneity. Whilst we condition on a range of characteristics, there may be factors that we cannot observe and yet may determine participation in E2E courses as well as labour market outcomes. We can see from the summary statistics reported in Table 1 that E2E participants are negatively selected. They come from a more disadvantaged background and performed more poorly at school exams. Assuming that E2E participants are also negatively selected on unobserved factors, then the OLS estimates of  $\beta$  are likely to be downward biased.

To overcome this endogeneity issue, we exploit geographical variation in the availability of E2E across Local Authority Districts (LADs). Figure A.1 in Appendix shows the proportion of

---

<sup>8</sup>Estimation was conducted in R using the `glmnet` package for the LASSO and the `xgboost` package for the gradient boosted trees.

providers offering E2E courses in each LAD in England. As can be seen from the map, there is substantial variation: E2E courses are not offered by any provider in 37 out of the 326 Local Authority Districts (LADs), while in some other LADs E2E courses are offered by over a third of providers. In the sample, the average proportion of providers offering E2E in the LAD is about 20 per cent. Figure A.2 shows the distribution of the proportion of providers offering E2E courses in the LADs, estimated for our sample of young people who stopped education after secondary school or participated in E2E. We can see that about five per cent of them live in LADs where there is no E2E provision.

In the first stage of this Instrumental Variable (IV) approach, we model E2E participation as a function of the proportion of providers offering E2E in the LAD where young people lived in their last year of secondary school and a range of local area and individual characteristics:

$$E2E_{i,j} = \delta propE2E_j + x_{i,j}\gamma + \eta_i \quad (3)$$

Where  $propE2E_j$  is the proportion of providers in the individuals home LAD  $j$  which offer E2E.  $x_{i,j}$  is a vector of relevant local and individual characteristics (we discuss below in more detail the variables we include). Because the instrument only varies at the LAD level, standard errors are clustered at the LAD level.

To be valid, the instrumental variable should satisfy three conditions. First, it should be correlated with the endogenous variable. This assumption can be easily tested.

Second, whilst the instrument may have no effect on some individuals, it must affect everybody who is affected in the same way (monotonicity). It is highly unlikely that the same young people who participate in E2E if not many providers offer E2E in their LAD would choose not to participate in E2E if it was offered by more providers in their LAD.

Third, the instrument should not be correlated with the error terms  $\eta_i$  and  $\varepsilon_i$ , which implies that the instrument must be as good as random, once we account for the covariates included in the model. Individuals who live in LADs in which many providers offer E2E should not differ in terms of unobserved characteristics ( $\eta_i$ ) from those living in LADs in which no or a few providers offer E2E courses. We can see from Figure A.1 that the LADs that do not offer

E2E courses are primarily located in the South East. In Figure A.3, we show the relationship between a range of socio-demographic characteristics of the young people in our sample and the proportion of providers which offer E2E courses in the LAD. Results suggest that LADs in which a higher proportion of providers offer the E2E programme tend to be more deprived compared to LADs in which few or no providers offer this programme. There is a strong relationship between the proportion of providers offering E2E courses and both the average IDACI Decile and the proportions eligible to Free School Meals or having special educational needs. We also find that LADs with high availability of E2E have on average lower achievement at KS3 and KS4 exams.

As living in LADs with low E2E availability is correlated with many observed background characteristics, omitting these factors in our model would lead the instrument to be correlated with the error terms ( $\eta_i$  and  $\varepsilon_i$ ) and could bias the estimates of the effect of E2E participation ( $\beta$ ).

To mitigate this issue, we add a range of covariates to our model. By doing so, the assumption that living in LADs that do not offer E2E courses is not correlated with the error term  $\eta_i$  is more likely to hold. Specifically, we include region-specific fixed effects, variables for deciles of Income Deprivation Affecting Children Index (IDACI) at LSOA level, the proportion of adults holding qualifications of various levels for each LSOA, and a measure of local unemployment in September 2003 (the number of claimants divided by working-age population). We also control for the number of FE providers in the LAD. Individual characteristics includes a range of socio-demographic factors (gender, special educational needs, and dummies for the white ethnic group and English as their first language). In our main specification, we also add secondary school attainment measured by the number of GCSEs obtained, and the number of GCSEs graded A\* -C and test results in English, Mathematics and Science administered at the age of 11 and 14.

We argue that conditional on these characteristics, the instrument can be considered as good as random. Although we control for a range of area-level characteristics, the proportion of providers that offer the E2E programme in the home LAD may still be correlated with unobserved area characteristics. We propose two placebo tests to assess the plausibility of this assumption.

First, we estimate the reduced-form effect of the instrument on the pre-determined educational outcomes (results at KS4, KS3 and KS2) of both E2E participants and those who left school at 16, conditioning on observed area-level characteristics. If the instrument is associated with educational outcomes before the age of 16, then it would indicate that the instrument acts as a proxy for unobserved area characteristics. As a result, the instrument would not be exogenous and therefore the 2SLS estimates would be biased. If there is no association, this test would provide some reassurance that the 2SLS estimates of the effect of E2E on labour market outcomes are not driven by unobserved area characteristics. This test assumes that area characteristics that affect educational outcomes also influence labour market outcomes. Whilst this assumption is reasonable, we propose another placebo test to assess whether the instrument picks up unobserved area characteristics which could influence labour market outcomes.

In our second set of placebo tests, we estimate the association of the instrument with the labour market outcomes of those who after leaving secondary school studied for A-levels, the upper-secondary qualifications which open the door to university. The proportion of providers that offer the E2E programme in the home LAD should not affect their labour market outcomes. Young people who studied for A-levels were not affected in any way by the provision of the E2E programme, and therefore their labour market outcomes should be independent of the provision of E2E. It may nonetheless be associated with their labour outcomes because the instrument acts as a proxy for unobserved area characteristics. The absence of association between the instrument and their labour market outcomes would indicate that the instrument does pick up the effect of unobserved area characteristics.

## **5 Results**

### **5.1 Linear and partially linear models**

In Table 2 we present OLS and LASSO estimates of the association between E2E participation and labour market outcomes in the 2015 tax year, conditional on various control variables. In column one, we show the unconditional association. In the 2015 tax year, those who participated in E2E earned £1,096 less than those who did not do any further education after leaving

school. However, they were 10.8 percentage points more likely to be employed and on average worked 34.4 days more in the 2015 year. Those observed in employment had lower earnings, by about a third.

Adjusting for area characteristics such as region dummies, the Income Deprivation Affecting Children Index, and the local unemployment rate in 2003 (column 2) and socio-demographic characteristics such as gender, free school meal eligibility, ethnicity and special need status increases the estimated coefficients for all outcomes. However, participation in E2E remains negatively associated with total earnings in the 2015 tax year. Controlling for prior educational attainment, such as results at national examination at age 11, 14 and 16 changes the sign of this association. Conditional on area characteristics, socio-demographic characteristics and schooling attainment, participation in E2E is associated with a £218 increase in annual earnings in the 2015 tax year. Participation in E2E is associated with a 12 percentage points increase in employment rate and an additional 48 days in employment. However, even when controlling for these characteristics, E2E participation remains negatively associated with number of days worked and earnings among those who work.

In column 5 and 6 we present estimates of the effect of E2E obtained using the Double Debiased Machine Learning (DML) estimator proposed by Chernozhukov et al. (2018). In column 5, the DML estimator is based on models estimated via LASSO. In column 6, the DML estimator is based on models estimated with gradient boosted trees. In Column 5 are based on a LASSO, The estimates are close to those obtained from the linear models presented in column 4. Capturing interactions and non-linearities does not change the estimated magnitude of the association between E2E participation and labour market outcomes. This suggests that the linear models are not misspecified. interactions and non-linearity are unlikely to be a source substantially change the estimated magnitude of the association between E2E participation and labour market outcomes.

## **5.2 Effect of E2E availability on participation**

Whilst we condition on a range of characteristics, there may be factors that we cannot observe and yet may determine participation in E2E as well as labour market outcomes. As discussed previously, E2E participants are negatively selected on observable characteristics. Assuming

Table 2: OLS and LASSO estimates of the effect of E2E on labour market outcomes

	OLS			LASSO	Boosted Trees	
	(1)	(2)	(3)	(4)	(5)	(6)
A. Earnings in 2015 tax year						
E2E	-1,095.711*** (70.569)	-644.017*** (71.491)	-263.203*** (69.956)	218.051*** (69.218)	180.018*** (63.569)	209.019*** (63.646)
Observations	99,260	99,260	99,260	99,260	99,260	99,260
Adjusted R <sup>2</sup>	0.002	0.019	0.086	0.129	-	-
B. Employed in 2015 tax year						
E2E e2e	0.087*** (0.003)	0.097*** (0.003)	0.115*** (0.003)	0.121*** (0.003)	0.117*** (0.003)	0.116*** (0.003)
Observations	99,260	99,260	99,260	99,260	99,260	99,260
Adjusted R <sup>2</sup>	0.007	0.013	0.041	0.068	-	-
C. Days in employment in 2015 tax year (excl 0)						
E2E	-5.666*** (0.524)	-4.750*** (0.536)	-3.449*** (0.545)	-2.197*** (0.551)	-2.434*** (0.471)	-1.983*** (0.446)
Observations	58,882	58,882	58,882	58,882	58,882	58,882
Adjusted R <sup>2</sup>	0.002	0.004	0.007	0.013	0.013	0.013
D. Log of earnings in 2015 tax year						
E2E	-0.327*** (0.010)	-0.284*** (0.010)	-0.234*** (0.010)	-0.180*** (0.010)	-0.193*** (0.010)	-0.182*** (0.010)
Observations	56,750	56,750	56,750	56,750	56,750	56,750
Adjusted R <sup>2</sup>	0.019	0.030	0.078	0.116	0.116	0.116
Covariates	None	Area characte- ristics	+ demo- graphics	+ KS2, KS3 & KS4 results Results	All interacted	All

Note: Demographic characteristics include Gender, Special Need Status, Free School Meal eligibility, ethnicity, english as first language. Area characteristics include region and IDACI deciles dummies, local unemployment rates, local level of education (measured by percentage of local residents with qualifications of various levels). KS4 results include number of GCSEs and number of GCSE A\*-C included as dummies, and whether they obtained a GCSE A\*-C in Mathematics, English or Science. KS2 and KS3 results include binary variables indicating if the student obtained the expected level at age 11 and 14 in English, Mathematics and Science. Standard errors in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Source: NPD, ILR and HMRC data, KS4 leavers 2002/03

that participants are also negatively selected on unobserved factors, the OLS estimates are likely to be downward biased, thereby understating the improvement in labour market performance that resulted from participating in the programme.

To overcome this endogeneity issue, we exploit geographical variation in the availability of E2E and use the proportion of providers offering E2E in the Local Authority District (LAD) where the young people completed secondary school as an instrument for participation. Table 3 shows OLS estimates of the effect of availability of E2E courses in the young peoples home LAD on participation in E2E across five different specifications including different sets of

covariates. In the first column, we present results from a model without additional covariates, reflecting the unconditional association between availability and participation.

As discussed above, the low availability of E2E is found in the more affluent areas, with relatively lower deprivation and higher educational attainment on average. Conditioning on area and individual characteristics is crucial for the instrument to be as good as random. In the second column, we show results from a model that includes region fixed effects, dummies for IDACI deciles home LSOA, the educational profile of the home LSOA in 2001 and the unemployment rate in September 2003 at LAD level as defined above. In column 3 we add demographic characteristics (gender, special educational needs, white ethnic group and English as their first language). In column 4 we add secondary school attainment measured by the number of GCSEs obtained and those graded A\*-C. Finally, in column 5 we add results to tests in English, Mathematics and Science administered at Key Stage 3.

Table 3: First Stage results: the effect of E2E availability on programme participation

	(1)	(2)	(3)	(4)	(5)
Proportion of providers providing E2E	0.3877*** (0.0600)	0.3209*** (0.0568)	0.3090*** (0.0572)	0.3108*** (0.0572)	0.3121*** (0.0563)
F- Stat. (excl. inst.)	41.814	31.952	29.179	29.544	30.781
Observations	99260	99260	99260	99260	99260
Covariates	Region	+ area characte- ristics	+ demo- graphics	+ KS4 results	+ KS2 and KS3 results

Note: Controls: Area characteristics include region and IDACI deciles dummies, local unemployment rates, local level of education (measured by percentage of local residents with qualifications of various levels) and number of Further education providers. Demographic characteristics include Gender, Special Need Status, ethnicity, English as first language. KS4 results include number of GCSEs and number of GCSE A\*-C included as dummies, and whether they obtained a GCSE A\*-C in Mathematics, English or Science. KS2 and KS3 results include binary variables indicating if the student obtained the expected level at age 11 and 14 in English, Mathematics and Science. Standard errors clustered at LAD level in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Source: NPD, ILR and HMRC data, KS4 leavers 2002/03

The most comprehensive and preferred specification (column 5) shows that a ten percentage points increase in the proportion of providers offering the programme increases the probability to participate in E2E by 3.1 percentage points, conditional on all other covariates. The F-statistic of 30.8 indicates that it is a strong predictor of E2E participation. As can be seen in Table 3, the estimates are robust to a change in the covariates included in the model.

### 5.3 Instrumental Variables estimates of the effect of E2E

In Table 4 we show Instrumental Variables estimates (two-stage least squares, 2SLS) of the effect of participating in E2E courses on earnings in the 2015/16 tax year. The first stage of this model uses the specification of column 5 of Table 3. We discuss the robustness of the estimates further below.

The main result indicates that participating in E2E resulted in a substantial increase in annual earnings (£7,750) in the 2015/16 tax year. This effect is large since the average annual earnings (including those with zero earnings) for the control group is £8,570.

This large increase in annual earnings can be due to an increase in the probability to be in employment, fewer days spent out of work during the year or higher wages. The data do not provide information on hourly or monthly wages. However, we estimate the effect of E2E on employment rate (defined as having done any paid work at some point in the tax year), the number of days employed in the tax year and the log of annual earnings, conditional on having been employed at some point during the year. We find that the increase in annual earnings is primarily driven by an increase in the probability to be in employment. Participating in E2E resulted in a 37 percentage points increase the probability to be employed in 2015/16 tax year. We find no statistically significant effect on employment duration during the tax year or on earnings conditional on being employed. The programme increased earnings by increasing the probability of being in work but not by increasing the employment duration nor increasing wages.

In Figure 2, we report the 2SLS estimates of the effect of participating in E2E courses on labour market outcomes in every tax year from 2003/04 to 2015/16, using the same model as used in Table 4.

We find that E2E raised annual earnings and increased the probability to be in employment. The estimates of the effects of E2E on annual earnings are shown in Graph A and indicate that the E2E programme significantly increased annual earnings in every tax year from 2008/09 onwards. Participating in E2E had no effect on earnings in 2003/04 and 2004/05 because most E2E participants may still have been enrolled in the E2E programme or in other vocational courses. The coefficients for the 2006/07 and 2007/08 tax years are positive but not statistically

Table 4: 2SLS estimates of effect of E2E on labour market outcomes in 2015

	(1) Earnings	(2) Employed	(3) Days employed	(4) log earnings (excl 0)
E2E	7750.491*** (2376.003)	0.372*** (0.097)	3.379 (10.501)	0.229 (0.215)
F- Stat. (excl. inst.)	30.781	30.781	28.022	25.361
Mean	8570.603	0.563	350.132	9.289
Observations	99,260	99,260	58,882	56,750

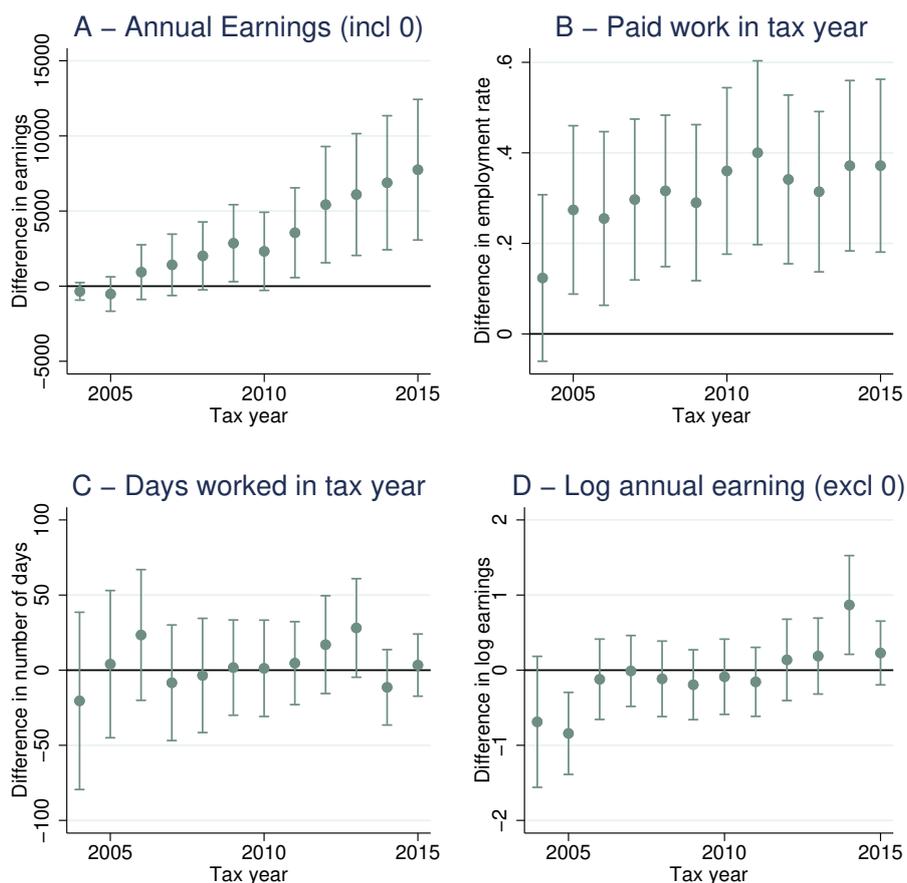
Note: Area characteristics include region and IDACI deciles dummies, local unemployment rates, local level of education (measured by percentage of local residents with qualifications of various levels) and number of Further education providers. Demographic characteristics include Gender, Special Need Status, ethnicity, English as first language. Prior educational attainment is measured by KS2, KS3 and KS4 results. KS4 results include number of GCSEs and number of GCSE A\*-C included as dummies, and whether they obtained a GCSE A\*-C in Mathematics, English or Science. KS2 and KS3 results include binary variables indicating if the student obtained the expected level at age 11 and 14 in English, Mathematics and Science. Standard errors clustered at LAD level in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Source: NPD, ILR and HMRC data, KS4 leavers 2002/03

significant. However, we can see that the returns to the programme have been increasing over time, from £2,500 to £7,750 in the 2015/16 tax year. Graph B shows that E2E significantly increased the probability to have done in paid work in every tax year but 2003/04, the tax year during which the majority had been enrolled in the programme. In all following tax years, E2E participation increased the employment rate by 30 to 40 percentage points.

However, E2E participation had little effect on time spent in employment or on pay amongst those who are employed. As shown in Graph C, participating in E2E had no effect on the number of days worked in the tax year, conditional on being having been employed at some point during the tax year. Similarly, Graph D shows that E2E participation did not significantly affect the log of annual earnings (excluding those who did not work). The increase in total earnings (including those who did not work) is primarily driven by the effect of the programme on their employment rate.

Figure 2: 2SLS estimate of effect of participating in E2E on labour market outcomes between 2003 and 2015



Note: Area characteristics include region and IDAC1 deciles dummies, local unemployment rates, local level of education (measured by percentage of local residents with qualifications of various levels) and number of Further education providers. Demographic characteristics include Gender, Special Need Status, ethnicity, English as first language. Prior educational attainment is measured by KS2, KS3 and KS4 results. KS4 results include number of GCSEs and number of GCSE A\*-C included as dummies, and whether they obtained a GCSE A\*-C in Mathematics, English or Science. KS2 and KS3 results include binary variables indicating if the student obtained the expected level at age 11 and 14 in English, Mathematics and Science. 95% confidence intervals.

Source: NPD, ILR and HMRC data, KS4 leavers

We further examine whether E2E participation affected socio-demographic groups differently. Table 5 reports 2SLS estimates of the effect of E2E participation on annual earnings in the 2015 tax year by gender, special educational needs status, free school meal eligibility and GCSE attainment. E2E participation had a positive and statistically significant on all subgroup, except for those who had been eligible for free school meals in secondary school. There appears to be little variation across the groups in terms of the magnitude of the effects relative to the control mean for each group, suggesting that the effects of the programme are rather homogeneous.

Table 5: 2SLS estimates of effect of E2E on earnings in 2015 by groups

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Women	Men	SEN	No SEN	FSM	no FSM	1+ GCSE A*C	No GCSE A*C
E2E	5543.0*** (2017.1)	9996.9*** (3589.1)	6923.7*** (2659.6)	7936.1*** (2727.1)	3965.0 (2619.7)	9262.0*** (2594.6)	11970.9*** (3939.6)	5424.8** (2251.1)
F- Stat. (excl. inst.)	29.873	18.731	22.207	29.143	24.695	29.502	25.341	29.621
Mean	5714.3	10672.9	6637.8	9326.2	5915.5	9302.1	10805.2	6888.4
Observations	40981	58279	33367	65893	24912	74348	36430	62830

Note: Outcome: total earnings in the 2015/16 tax year, including zeros. Instrument: proportion of providers offering E2E in home LAD. Controls: Area characteristics include region and IDACI deciles dummies, local unemployment rates, local level of education (measured by percentage of local residents with qualifications of various levels) and number of Further education providers. Demographic characteristics include Gender, Special Need Status, ethnicity, English as first language. Prior educational attainment is measured by KS2, KS3 and KS4 results. KS4 results include number of GCSEs and number of GCSE A\*-C included as dummies, and whether they obtained a GCSE A\*-C in Mathematics, English or Science. KS2 and KS3 results include binary variables indicating if the student obtained the expected level at age 11 and 14 in English, Mathematics and Science. Standard errors clustered at LAD level in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1  
Source: NPD, ILR and HMRC data, KS4 leavers 2002/03

A substantial share of E2E participants achieved a vocational qualification at Level 2 or above after completing the programme. We investigate whether the effect of the E2E programme is mediated by the attainment of a Level 2+ vocational qualification. We test whether the effect of E2E differs depending on whether the participants achieved a vocational qualification, by splitting the E2E participants into two groups depending on whether they achieved a Level 2+ qualification after E2E. We then estimate the effect of E2E for each group, using young people with no post-16 as a control group. As for our main results, we instrument E2E participation using the availability of E2E in the home LAD.

Results reported in Table 6 indicate that E2E had a positive effect on annual earnings and the probability of being employed for both those who achieved a Level 2 or above vocational qualification after E2E or those who did not. The estimate of the effect of E2E is lower for those who did not obtain a Level 2 or above qualification after E2E than for those who did. The results nonetheless indicate that the E2E programme improved young people's labour outcomes

regardless of whether they completed an additional vocational course.

E2E improved long-term labour market outcomes not only by increasing the probability to attain a Level 2 or higher qualification but is likely to have had a direct effect on labour market outcomes. The focus on non-cognitive skills, such as building confidence, managing own feelings and working in a team, may explain why the programme improved the labour market outcomes of those who did not engage further in vocational education and obtain a recognised qualification.

Table 6: Impact of E2E: the role of qualification attainment

	(1) Earnings	(2) Employed	(3) Days employed (excl 0)	(4) log earnings (excl 0)
A. No Level 2+ qualification achieved				
E2E	8,486.5*** (3045.3)	0.415*** (0.129)	3.2 (13.7)	0.270 (0.279)
F- Stat. (excl. inst.)	28.804	28.804	26.143	22.534
Mean	7,945.5	0.575	347.8	9.166
Observations	87,959	87,959	50,576	48,624
B. Achieved a Level 2+ qualification				
E2E	13,293.8*** (3436.8)	0.504*** (0.140)	14.9 (13.3)	0.695** (0.298)
F- Stat. (excl. inst.)	26.159	26.159	23.722	22.001
Mean	8,793.5	0.589	349.9	9.267
Observations	75,803	75,803	44,612	43,741

Note: Labour market outcomes in 2015. Control group: Young people with no post-16 education within three years after leaving school. Controls: Area characteristics (region and IDACI deciles dummies, local unemployment rates, local level of education - measured by percentage of local residents with qualifications of given levels) and number of Further education providers.), demographic characteristics (Gender, Special Need Status, ethnicity, English as first language) and KS2, KS3 and KS4 results. KS4 results include number of GCSEs and number of GCSE A\*-C included as dummies, and whether they obtained a GCSE A\*-C in Mathematics, English or Science. KS2 and KS3 results include binary variables indicating if the student obtained the expected level at age 11 and 14 in English, Mathematics and Science. Standard errors clustered at LAD level in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Source: NPD, ILR and HMRC data, KS4 leavers 2002/03

## 5.4 Sensitivity Analysis

### Different specifications

We show results from IV models based on different specification in Table B.1 in Appendix B. As discussed in the method section, we expect that the IV coefficients would be downward biased if we did not account for the fact that young people living in LADs with no or few E2E providers tend to be positively selected. This is what we find in column 1: the estimate of the model that with no additional covariates is negative for annual earnings including and excluding zeros earnings. The estimates are positive but not statistically significant for employment outcomes. Adjusting for regions, the ADICI, the educational profile of the home LSOA in 2001 and the local unemployment rate in 2003 substantially reduces the magnitude of this downward bias (column 2). However, controlling for individual-level characteristics, such as socio-demographic characteristics (column 3) and prior attainments (column 4) has little effect on the IV estimates. Column 4 is our preferred specification and is used in Figure 2. In column 5, we report results from a model using the log of the proportion of providers that offer the E2E programme in the home LAD instead of using the variable linearly. We find that the estimate of the effect E2E is in line with the results from our preferred instrument. In columns 6 and 7 we show results obtained with different control groups. In column 6, we use young people who engage in low-level vocational education ("Below Level 2") other than E2E courses as control group. The estimated effect of participating in E2E is lower but still large. In column 7, we show results from a model where the control group is restricted to those with no post-16 education and did not work in the first year after leaving school. Changing the control group does not substantially affect the results. Restricting our sample to individuals who were retrieved in the HMRC database does not alter our results either (column 8). Finally, we show that our results are robust to not clustering standard errors at the LAD level (column 9).

### Placebo tests

As discussed above, although we control for a wide range of observed characteristics, the instrument (i.e. the proportion of providers that offer the E2E programme) in the home LAD may still be correlated with unobserved variables part of the error term. Because the instru-

ment is positively associated with observed factors that are likely to result in lower earnings (such as poor educational attainment), it is also likely to be positively associated with unobserved factors that drive lower earnings. Those living in LADs with high E2E availability are likely to have unobserved characteristics that are associated with lower earnings, in which case the IV estimates are likely to be downward biased, resulting in underestimating the effects of participating in E2E on earnings.

As outlined in Section 4, we conduct two placebo tests to check whether the instrument is capturing the effect of unobserved characteristics. First, we estimate the reduced-form effect of the instrument on educational attainment at age 16 (KS4), 14 (KS3) and 11 (KS2). If the instrument acts as a proxy for unobserved area characteristics, we could expect to find differences that the estimates of reduced-form effect on educational attainment goes in the same direction as for earnings.

Table 7: Placebo test I: Reduced-Form effect of the instrument on educational attainment at age 11, 14 and 16

	(1)	(2)	(3)	(4)	(5)	(6)
	KS4		KS3		KS2	
	Number of GCSEs	Number of GCSEs A*-C	Level 5+ in English	Level 5+ in Maths	Level 4+ in English	Level 4+ in Maths
Prop. providing E2E	-0.034 (0.342)	0.024 (0.146)	0.024 (0.039)	0.008 (0.033)	0.006 (0.034)	-0.003 (0.035)
Mean	5.656	1.209	0.333	0.377	0.451	0.463
R-Squared	0.193	0.139	0.162	0.139	0.107	0.084
Observations	99,260	99,260	99,260	99,260	99,260	99,260

Note: KS4 examinations are taken at age 16, KS3 at 14 and KS2 at 11. Controls: Area characteristics (region and IDACI deciles dummies, local unemployment rates, local level of education - measured by percentage of local residents with qualifications of given levels) and number of Further education providers.), demographic characteristics (Gender, Special Need Status, ethnicity, English as first language). Standard errors clustered at LAD level in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1  
Source: NPD, ILR and HMRC data, KS4 leavers 2002/03

In Table 7, we show estimates of the reduced-form effect of the instrument on performance at KS4, KS3 and KS2 examinations. To measure performance at KS4, we use the number of GCSEs obtained, as well as the number of GCSEs graded A\*-C. For KS3, we use binary variables indicating if the young person obtained level 5 or over (the level expected at this age) in English and Mathematics. For KS2, we use binary variables indicating if the young person obtained level 4 or over (the level expected at this age) in English and Mathematics. We find that there is no association between the instrument, the proportion of providers offering E2E, and educational attainment, once we adjust for area and socio-demographic characteristics.

The estimates are close to zero and not statistically significant. This suggests that the effect on labour market outcomes are unlikely to be driven by unobserved area characteristics, providing some evidence that the exclusion restriction is unlikely to be violated.

Second, we estimate the association of the instrument with the labour market outcomes of those who studied for A-levels after leaving secondary school. The proportion of providers that offer the E2E programme in the home LAD should not have an effect on their labour market outcomes unless the instrument acts as a proxy for unobserved area characteristics.

In Table 8, we show the estimates of the association between the proportion of providers that offer the E2E programme in the home LAD and the labour market outcomes of those who studied for A-levels, conditional on the same variables as for our main IV models. We find no evidence that the instrument is associated with the labour market outcomes of those who studied for A-levels. This provides some reassurance that the instrument is not a proxy for unobserved areas characteristics, and therefore that the exclusion restriction is not violated.

Table 8: Placebo test II: Reduced-form effect of instrument on the labour market outcomes of those who studied for A-Levels

	(1) Earnings	(2) Employed	(3) Days employed (excl 0)	(4) log earnings (excl 0)
Proportion of providers offering E2E	392.9 (769.4)	0.016 (0.010)	5.228 (3.969)	0.021 (0.038)
Mean	18,983.6	0.836	357.624	9.791
R-Squared	0.044	0.008	0.008	0.045
Observations	255,577	255,577	213,703	212,052

Note: Sample restricted to young people who left school in 2003 and went on to study for A-levels. Controls: Area characteristics (region and IDACI deciles dummies, local unemployment rates, local level of education - measured by percentage of local residents with qualifications of given levels) and number of Further education providers.), demographic characteristics (Gender, Special Need Status, ethnicity, English as first language) and KS2, KS3 and KS4 results. KS4 results include number of GCSEs and number of GCSE A\*-C included as dummies, and whether they obtained a GCSE A\*-C in Mathematics, English or Science. KS2 and KS3 results include binary variables indicating if the student obtained the expected level at age 11 and 14 in English, Mathematics and Science. Standard errors clustered at LAD level in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1  
Source: NPD, ILR and HMRC data, KS4 leavers 2002/03

## 6 Conclusion

We estimate the long-term effects on earnings of the Entry to Employment (E2E) programme, an intervention designed to improve the non-cognitive skills of low-achieving adolescents in England. We exploit the substantial heterogeneity in the availability of E2E courses across Local Authority Districts to overcome the issue of selection into the programme. We find that the availability of E2E is a strong predictor of E2E participation, even when controlling for LAD and individual characteristics. Our results indicate that the programme substantially increased the earnings of the participants in the medium and longer term. We find that E2E participation increased earnings by over £7,000 in the 2015/16 tax year. This increase in annual earnings is largely driven by an increase in the probability to be employed. We find no evidence that E2E increased pay or number of days worked per year amongst those who are employed. The improvement in labour market outcomes does not only result from E2E allowing young people to pursue vocational courses after completing the programme. We find that E2E raised the earnings of those who did not complete a vocational course after E2E, suggesting that this programme focusing on non-cognitive skills had a direct effect on long-term labour market outcomes.

We contribute to the existing literature by showing that interventions aiming to improve the non-cognitive skills of low-achieving adolescents can substantially improve their labour market outcomes in the longer term. During the adolescent years, non-cognitive skills are more malleable than cognitive skills (Kautz et al. 2014), and therefore interventions should focus on soft skills. Our results also highlight the importance of focusing on long-term outcomes. If we had only examined the short-term effects of the programme, we would have erroneously concluded that the programme did not affect the participants' labour market outcomes. This lack of long-term follow-ups for programmes that target adolescents may have led researchers to wrongly conclude that these programmes have little effect on the lives of the participants.

## References

- Almlund, M., Duckworth, A. L., Heckman, J. & Kautz, T. (2011), 'Personality Psychology and Economics', in: *E.Hanushek, S.Machin and L.Woessman (eds.), Handbook of the Economics of Education* pp. 1–81.
- Borghans, L., Duckworth, A. L., Heckman, J. J. & ter Weel, B. (2008), 'The Economics and Psychology of Personality Traits', *Journal of Human Resources* **43**(4), 972–1059.
- Borghans, L., Weel, B. t. & Weinberg, B. A. (2008), 'Interpersonal styles and labor market outcomes', *Journal of Human Resources* **43**(4), 815–858.
- Caliendo, M., Cobb-Clark, D. & Uhlenhorff, A. (2015), 'Locus of Control and Job Search Strategies', *Review of Economics and Statistics* **97**(1), 88–103.
- Carneiro, P., Crawford, C. & Goodman, A. (2007), The impact of early cognitive and non-cognitive skills on later outcomes, CEE discussion papers number 92, Centre for the Economics of Education, LSE.
- Chen, T. & Guestrin, C. (2016), Xgboost: A scalable tree boosting system. in 'Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining', ACM, pp. 785–794.
- Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W. & Robins, J. (2018), 'Double/debiased machine learning for treatment and structural parameters', *The Econometrics Journal* **21**(1), C1–C68.
- Chetty, R., Friedman, J. N., Hilger, N., Saez, E., Schanzenbach, D. W. & Yagan, D. (2011), 'How does your kindergarten classroom affect your earnings? Evidence from project star', *Quarterly Journal of Economics* .
- Cook, P. J., Dodge, K., Farkas, G., Fryer, Jr, R. G., Guryan, J., Ludwig, J., Mayer, S., Pollack, H. & Steinberg, L. (2014), 'The (Surprising) Efficacy of Academic and Behavioral Intervention with Disadvantaged Youth: Results from a Randomized Experiment in Chicago', *National Bureau of Economic Research* pp. 1–59.
- Deere, A. (2016), Progression or cul-de-sac? From foundation learning to the stud programme, A study of four organisations in the English Further Education Sector, PhD thesis, Institute of Education, University College London.
- Department for Education and Skills (2001), Modern Apprenticeship: The Way to Work, Technical report, Report of the Modern Apprenticeship Advisory Committee.
- Dur, R., Non, A. & Roelfsema, H. (2010), 'Reciprocity and incentive pay in the workplace', *Journal of Economic Psychology* **31**(4), 676–686.
- Feinstein, L. (2000), The relative economic importance of academic, psychological and behavioural attributes developed on childhood, CEE discussion papers number 443, Centre for the Economics of Education, LSE.
- Gallo, W., Bradley, E. & Kasl, S. (2003), 'The influence of internal control on the employment status of german workers', *Schmoller Jahrbuch* **123**(1), 1–15.
- Heckman, J. J. (2006), 'Skill formation and the economics of investing in disadvantaged children', *Science* **312**(5782), 1900–1902.
- Heckman, J. J. & Kautz, T. (2012), 'Hard evidence on soft skills', *Labour Economics* **19**(4), 451–464.

- Heckman, J. J., Moon, S. H., Pinto, R., Savelyev, P. A. & Yavitz, A. (2010), 'The rate of return to the HighScope Perry Preschool Program', *Journal of Public Economics* **94**(1-2), 114–128.
- Heller, S. B., Shah, A. K., Guryan, J., Ludwig, J., Mullainathan, S. & Pollack, H. A. (2017), 'Thinking, fast and slow? Some field experiments to reduce crime and dropout in Chicago', *Quarterly Journal of Economics* **132**(1), 1–54.
- Holmlund, H. & Silva, O. (2014), 'Targeting noncognitive skills to improve cognitive outcomes: Evidence from a remedial education intervention', *Journal of Human Capital* **8**(2), 126–160.
- Jacob, B. A. & Lefgren, L. (2004), 'Remedial Education and Student Achievement: A Regression-Discontinuity Analysis', *Review of Economics and Statistics* **86**(1), 226–244.
- Kautz, T., Heckman, J. J., Diris, R., ter Weel, B. & Borghans, L. (2014), 'Fostering and Measuring Skills: Improving Cognitive and Non-Cognitive Skills to Promote Lifetime Success', *IZA Discussion Papers* **No. 8696**.
- Lavy, V. & Schlosser, A. (2005), 'Targeted remedial education for underperforming teenagers: Costs and benefits', *Journal of Labor Economics* **23**(4), 839–874.
- Learning and Skills Council (2009), *Making Skills Matter; Learning and Skills Council's Annual Report and Accounts for 200809*, The Stationery Office, London.
- Lindqvist, E. & Vestman, R. (2011), 'The Labor Market Returns to Cognitive and Noncognitive Ability: Evidence from the Swedish Enlistment', *American Economic Journal: Applied Economics* **3**(1), 101–128.
- Machin, S., Mcintosh, S., Vignoles, A. & Viitanen, T. (2001), *Basic Skills, Soft Skills and Labour Market Outcomes*, DfE research centre brief 250, UK Department for Education.
- Martins, P. S. (2017), (How) Do Non-Cognitive Skills Programs Improve Adolescent School Achievement? Experimental Evidence, IZA Discussion Papers 10950, Institute of Labor Economics (IZA).
- Rodríguez-Planas, N. (2012), 'Longer-term impacts of mentoring, educational services, and learning incentives', *American Economic Journal: Applied Economics* **4**(4), 121–139.
- Schochet, P. Z., Burghardt, J. & McConnell, S. (2008), 'Does job corps work? impact findings from the national job corps study', *American Economic Review* **98**(5), 1864–86.
- Select Committee on Education and Skills (2003), *Entry to Employment (E2E), Memorandum submitted by the Learning and Skills Council*.
- Spielhofer, T., Mann, P. & Sims, D. (2003), *Entry to Employment (E2E) Participant Study, Final Report*, Slough, UK, National Foundation for Educational Research, NFER.
- Strmer, S. & Fahr, R. (2013), 'Individual determinants of work attendance: evidence on the role of personality', *Applied Economics* **45**(19), 2863–2875.
- Taylor, E. (2014), 'Spending more of the school day in math class: Evidence from a regression discontinuity in middle school', *Journal of Public Economics* **117**, 162–181.
- Wolf, A. (2011), 'Wolf Review of Vocational Education', *Department for Education, UK government* .

# Appendices

## A Availability of E2E

Figure A.1: Proportion of providers offer E2E by Local Authority District in the 2003/04 academic year

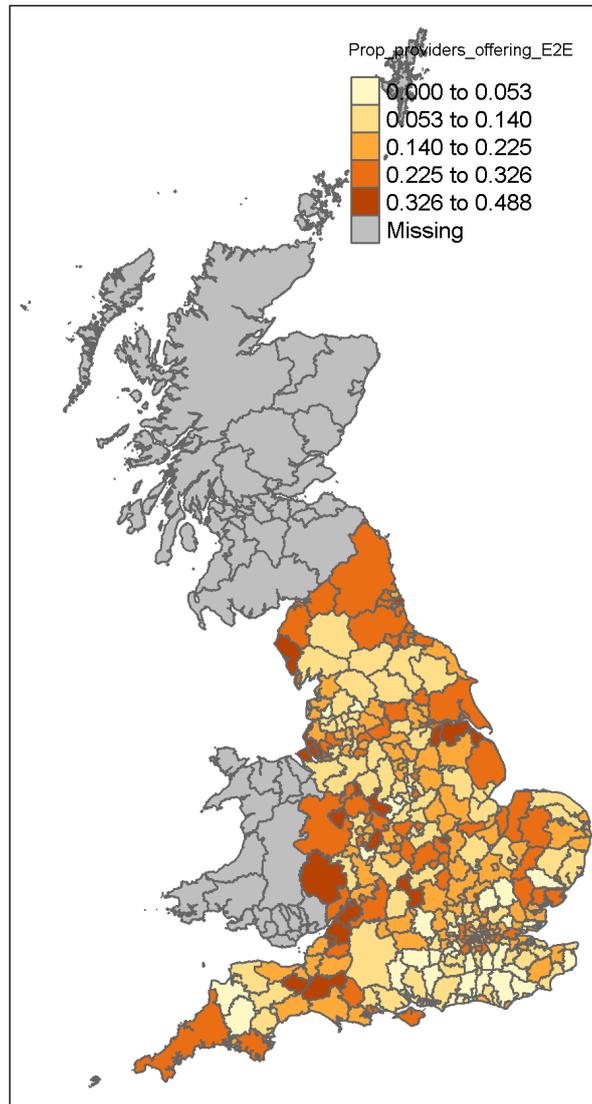


Figure A.2: Distribution at the individual level of the proportion of providers offering E2E at the LAD level

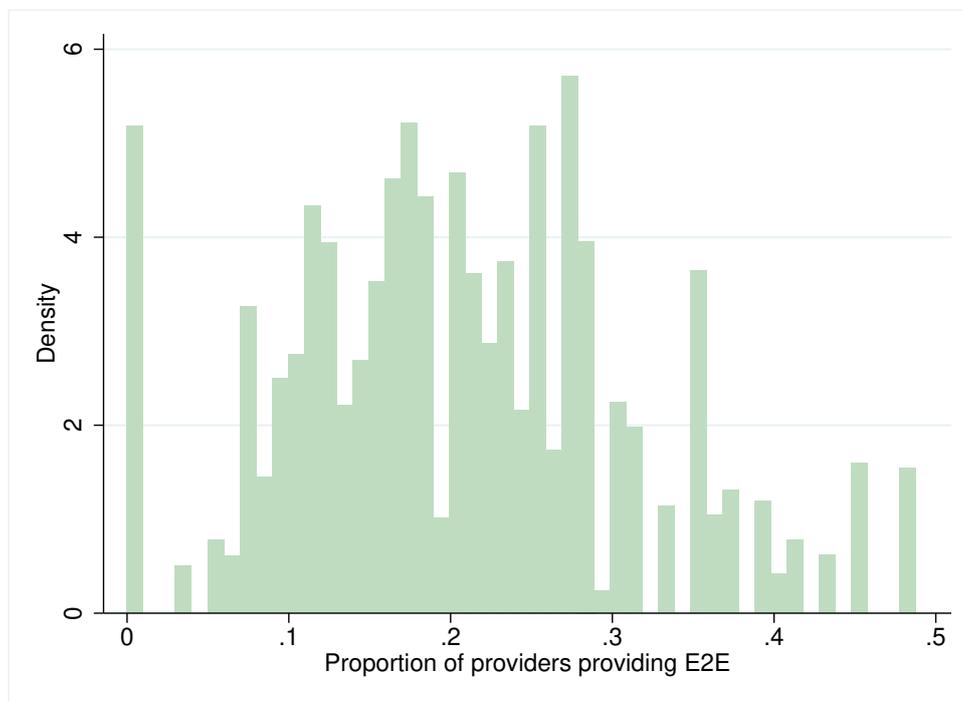
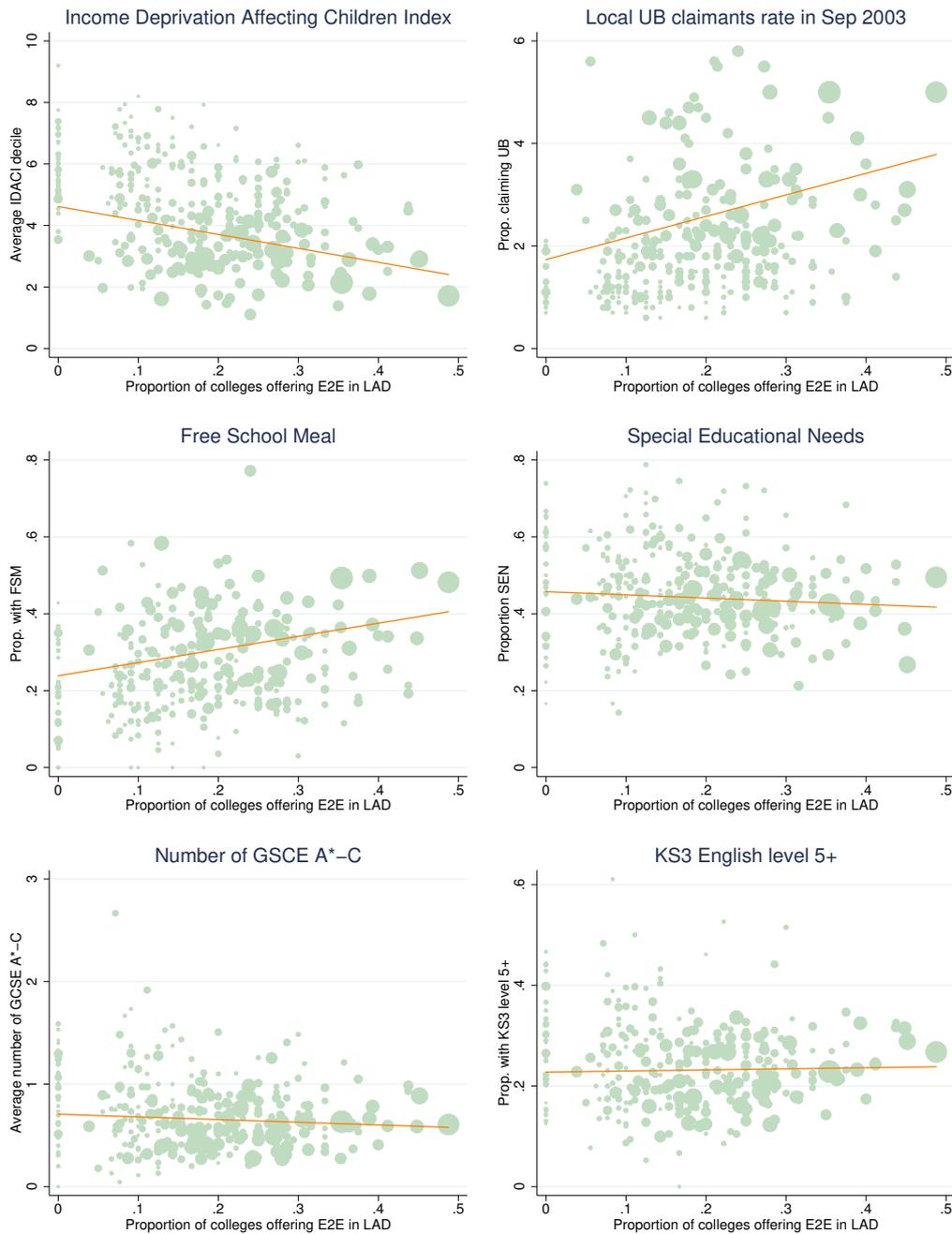


Figure A.3: Association between socio-demographic characteristics and proportion of colleges offering E2E in the LAD



## B Sensitivity analysis

Table B.1: 2SLS estimates of the effects of E2E participation on earnings in 2015/16: different specifications

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A. Earnings in 2015 tax year									
E2E	-2570.5*	7748.2***	8043.3***	7750.5***	7693.3***	3840.6*	7931.8***	6710.0***	7750.5***
	(1367.1)	(2515.0)	(2714.5)	(2376.0)	(2460.8)	(2091.5)	(1237.9)	(2516.6)	(1212.8)
F- Stat. (excl. inst.)	68.312	31.952	29.179	30.781	19.007	26.737	266.598	30.503	362.984
Observations	99260	99260	99260	99260	99260	99511	71830	82554	99260
B. Employed in 2015 tax year									
E2E	0.062	0.377***	0.386***	0.372***	0.361***	0.163**	0.486***	0.261***	0.372***
	(0.042)	(0.093)	(0.105)	(0.097)	(0.094)	(0.073)	(0.062)	(0.093)	(0.056)
F- Stat. (excl. inst.)	68.312	31.952	29.179	30.781	19.007	26.737	266.598	30.503	362.984
Observations	99260	99260	99260	99260	99260	99511	71830	82554	99260
C. Days employed in 2015 tax year (excl 0)									
E2E	-11.46**	3.10	3.77	3.38	5.59	-3.54	-0.96	0.78	3.38
	(5.79)	(11.12)	(11.38)	(10.50)	(10.65)	(9.43)	(12.42)	(10.65)	(8.74)
F- Stat. (excl. inst.)	58.750	27.605	26.671	28.022	16.804	26.634	132.021	28.174	235.285
Observations	58882	58882	58882	58882	58882	66752	37529	58174	58882
D. log of earnings in 2015 tax year (excl 0)									
E2E	-0.567***	0.207	0.271	0.229	0.341	-0.191	0.145	0.229	0.229
	(0.153)	(0.258)	(0.256)	(0.215)	(0.242)	(0.188)	(0.221)	(0.215)	(0.168)
F- Stat. (excl. inst.)	54.842	24.863	24.107	25.361	16.347	25.879	119.558	25.361	203.819
Observations	56750	56750	56750	56750	56750	65106	36971	56750	56750
Instrument		% provider offering E2E in LAD			log of prop. in LAD		% provider offering E2E in LAD		
Covariates	None	Region,IMD deciles local UE	+ demo-graphics	+ KS3 and KS4 results	Same as (4)	Same as (4)	Same as (4)	Same as (4)	Same as (4)
Control group			No FE			Any BL2	No FE & not working in 2004	No FE	
								Exclude unmatched HMRC	Unclustered SE

Note: demographic characteristics include Gender, Special Need Status, Free School Meal eligibility, ethnicity, English as first language. KS4 results include number of GCSEs and number of GCSE A\*-C included as dummies, and whether they obtained a GCSE A\*-C in Mathematics, English or Science. KS2 and KS3 results include binary variables indicating if the student obtained the expected level at age 11 and 14 in English, Mathematics and Science. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Source: NPD, ILR and HMRC data, KS4 leavers 2002/03